### REMARKS

Applicants have substituted into the present specification a new paper copy Sequence Listing section according to 37 C.F.R. §1.821(c) as new pages 1-8.

Applicants have amended the previously filed Sequence Listing and revised Figures 1, 2A, 5, 7 and 8 to make them conform to the amended Sequence Listing.

SEQ ID NO:1 of the previously filed Sequence Listing corresponded to originally filed Figure 1. Originally filed Figure 1 erroneously duplicated the lower nucleotide sequence presented in Figure 4 (which lower sequence now corresponds to SEQ ID NO:4 in the attached substitute Sequence Listing). SEQ ID NO:1 of the Sequence Listing and Figure 1 have been amended to conform to the nucleotide sequence shown Figs. 2A (lower sequence), 5, 7 and 8. These amendments are supported by the disclosure in the specification of Figs. 1, 2A, 5, 7 and 8, that SEQ ID NO:1 is identical in all these figures.

The amendment to SEQ ID NO:2 was necessitated due to an erroneous addition of incorrect nucleotides at positions 886-953. These incorrect nucleotides have now been deleted from SEQ ID NO:2 and this amendment is supported by the nucleotide sequence shown in Figure 2A (upper sequence).

SEQ ID NO:3 has been deleted from the previously filed Sequence Listing as SEQ ID NO:3 was a partial sequence of SEQ ID NO:1. The subsequent sequences have been renumbered accordingly in the attached substitute Sequence Listing.

The revisions to Figs. 1, 2A and 5 were necessitated due to typographical/clerical errors made when drafting these figures, and were made to make the sequences of Figs. 1, 2A and 5 conform to SEQ ID NO:1.

The revisions to Figures 7 and 8 merely correct an alignment error in the second column of the last row of the sequence.

All of the above revisions to the Sequence Listing and the figures are of clerical/typographical nature and do not constitute new matter. Approval of all of the revisions is therefore respectfully requested.

Furthermore, attached hereto is a 3 1/2" disk containing the "Sequence Listing" in computer readable form in accordance with 37 C.F.R. §1.821(e).

Applicants have amended the specification to insert SEQ ID Nos, as supported in the present specification.

The following statement is provided to meet the requirements of 37 C.F.R. §1.825(a) and 1.825(b).

I hereby state, in accordance with 37 C.F.R. §1.825(a), that the amendments included in the substitute sheets of the sequence listing are believed to be supported in the application as filed and that the substitute sheets of the sequence listing are not believed to include new matter.

I hereby further state, in accordance with 37 C.F.R. §1.825(b), that the attached copy of the computer readable form is the same as the attached substitute paper copy of the sequence listing.

Under U.S. rules, each sequence must be classified in <213> as an "Artificial Sequence", a sequence of "Unknown" origin, or a sequence originating in a particular organism, identified by its scientific name.

Neither the rules nor the MPEP clarify the nature of the relationship which must exist between a listed sequence and an organism for that organism to be identified as the origin of the sequence under <213>.

Hence, counsel may choose to identify a listed sequence as associated with a particular organism even though that sequence does not occur in nature by itself in that organism (it may be, e.g., an epitopic fragment of a naturally occurring protein, or a cDNA of a naturally occurring mRNA, or even a substitution mutant of a naturally occurring sequence). Hence, the identification of an organism in <213> should not be construed as an admission that the sequence *per se* occurs in nature in said organism.

Similarly, designation of a sequence as "artificial" should not be construed as a representation that the sequence has no association with any organism. For example, a primer or probe may be designated as "artificial" even though it is necessarily complementary to some target sequence, which may occur in nature. Or an "artificial" sequence may be a substitution mutant of a natural sequence, or a chimera of two or more natural sequences, or a cDNA (i.e., intron-free sequence) corresponding to an intron-containing gene, or otherwise a fragment of a natural sequence.

The Examiner should be able to judge the relationship of the enumerated sequences to natural sequences by giving full consideration to the specification, the art cited therein, any further art cited in an IDS, and the results of his or her sequence search against a database containing known natural sequences.

Attached hereto is a marked-up version of the changes made to the specification by the current amendment. The attached page is captioned "Version with markings to show changes made".
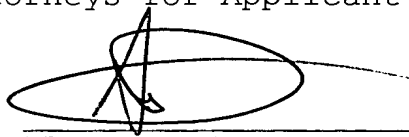
Applicants submit that the present application contains patentable subject matter and therefore urge the examiner to pass the case to issuance.

If the examiner has any questions or comments concerning the above described application, the examiner is urged to contact the undersigned at the phone number below.

Respectfully submitted,

BROWDY AND NEIMARK, P.L.L.C.
Attorneys for Applicant(s)

By _____
ALLEN C. YUN
Registration No. 37,971

ACY:al
624 Ninth Street, N.W.
Washington, D.C. 20001
Telephone No.: (202) 628-5197
Facsimile No.: (202) 737-3528
F:\,C\cohn\moroz3\pto\RESPONSE TO NOTICE TO COMPLY.doc

## VERSION WITH MARKINGS TO SHOW CHANGES MADE

**In the specification:**

The table on page 15 has been amended as follows:

| | |
|---|---|
| 5′ GGT GGC GAC GAC TCC TGG AGC CCG 3′ | SEQ ID NO:6 |
| 5′ TTG ACA CCA GAC CAA CTG GTA ATG 3′ | SEQ ID NO:7 |
| 5′ GAC CGC GAT GAT GTG GCT TTG AAG AAC 3′ | SEQ ID NO:8 |
| 5′ GAT AGG ATC TTT AGC GAC AGC CGA 3′ | SEQ ID NO:9 |
| 5′ ATG GCG GCC TCT GAG TCC TGG TGG 3′ | SEQ ID NO:10 |
| 5′ CGG GCT GAA TGC AAT GGA GTG TGC 3′ | SEQ ID NO:11 |
| 5′ GAC CCC CAT TTG TGT GAC 3′ | SEQ ID NO:12 |
| 5′ CGA CGA CTC CTG GAG CCC G 3′ | SEQ ID NO:13 |
| 5′ Biotin-TTG ACA CCA GAC CAA CTC GTA ATG 3′ | SEQ ID NO:14 |
| 5′ AGC CGA CAG CGA TTT CTA GGA TAG 3′ | SEQ ID NO:15 |
| 5′ GTT CTT CAA AGC CAC ATC ATC GCG GTC 3′ | SEQ ID NO:16 |
| 5′ GCT TTC ATT ATC ACT GTC TCC CAG GGT G 3′ | SEQ ID NO:17 |
| 5′ CAG ACG TTC TTC GCC GAG AGT CGT 3′ | SEQ ID NO:18 |
| 5′ CAG ACG TTC TTC GCC GAG AGT CGT CGG 3′ | SEQ ID NO:19 |
| 5′ CAT TTC GGG GAT TCG GGG GA 3′ | SEQ ID NO:20 |
| 5′ GGG GGA CGG AAC CCG GCG CT 3′ | SEQ ID NO:21 |
| 5′ CCC TCT ACA CTT ATC ATC TTC 3′ | SEQ ID NO:22 |
| 5′ CTA TCC TAG AAA TCG CTG TCG GCT 3′ | SEQ ID NO:23 |
| 5′ GTC ACT ACT GGA ATT CCC TTC TCC 3′ | SEQ ID NO:24 |
| 5′ GGA GAA GGG AAT TCC AGT AGT GAC 3′ | SEQ ID NO:25 |
| 5′ GGA AAT CGC TGT CGC CTA ACC 3′ | SEQ ID NO:26 |
| 5′ GGT TAG GCG ACA GCG ATT TCC 3′ | SEQ ID NO:27 |
| 5′ GGC CAC GCG TCG ACT AGT AC 3′ | SEQ ID NO:28 |
| 5′ GTA ATG CAC ACTCCA TTG GC 3′ | SEQ ID NO:29 |
| 5′ GTA ATG CAC ACT CCA TTG 3′ | SEQ ID NO:30 |
| 5′ GCG CTC AGC TGG AAT TCC 3′ | SEQ ID NO:31 |
| 5′ GGA ATT CCA GCT GAG CGC 3′ | SEQ ID NO:32 |
| 5′ GTG GGA TCC CCA TGA CGA CCG CGT CCA CC 3′ | SEQ ID NO:33 |
| 5′ GAC TCG AGT TAA GCC GAC AGC GAT TTC 3′ | SEQ ID NO:34 |
| 5′ GAC TCG AGT CAG GGT GAC CGA AAA ATC AG 3′ | SEQ ID NO:35 |
| 5′ CCC GCT CGA GTC AGG GTG ACC GAA AAA TCA G 3′ | SEQ ID NO:36 |

The two paragraphs beginning at line 6 on page 16 and ending at line 13 of page 16 have been amended as follows:

**Fig. 1** shows the nucleic acid sequence (SEQ ID NO:1) of clone T16 isolated from T47D breast cancer cDNA library. Initiation and termination codons of the open reading frame are indicated by dark bars;

**Fig. 2A** shows a comparison of the nucleic acid sequences (upper sequence) (SEQ ID NO:2) of clone 4.7 isolated from a placenta cDNA library exhibiting normal human FTH, and the sequences (lower sequence) of clone T16 (SEQ ID NO:1) isolated from human breast cancer T47D cDNA library. Initiation and termination codons of the open reading frame are marked by dark boxes;

The three paragraphs beginning at line 17 of page 16 and ending at line 25 of page 16 have been amended as follows:

**Fig. 3** shows a comparison of sequence homology between cDNA clone T16 (residues 463-671 of SEQ ID NO:1) and human mitochondrial cytochrone oxidase I DNA (SEQ ID NO:3);

**Fig. 4** shows a comparison of nucleic acid sequences between placental cDNA obtained by PCR amplification using T16 specific primers (upper sequence) (residues 24-822 of SEQ ID NO:1) and T16 cDNA sequence obtained from the T16 cDNA clone (lower sequence) (SEQ ID NO:4). Identical nucleic acid sequences are indicated by a dotted line. Initiation and termination codons are indicated by a dark bar;

**Fig. 5** shows the nucleic acid sequence and deduced amino acid sequence (SEQ ID NO:5) of the cDNA of OFF1;

The two paragraphs beginning at line 4 of page 17 and ending at line 6 of page 17 have been amended as follows:

**Fig.** 7 shows the sequence of clone T16 (SEQ ID NO:1). pPrimers used for PCR are indicated in the above sequence;

**Fig.** 8 shows the restriction enzyme map sequence of clone T16 (SEQ ID NO:1);

Table 1 on page 19 has been amended as follows:

**Table 1**

**List of Primers**

| Name | #MR | Sequence | SEQ ID NO: | | |
|------|-----|----------|------------|---|---|
| 1060F | 24 | 5′ GGT GGC GAC GAC TCC TGG AGC CCG 3′ | 6 | 75% | |
| 1061R | 24 | 5′ TTG ACA CCA GAC CAA CTG GTA ATG 3′ | 7 | 45.80% | |
| 17F | 27 | 5′ GAC CGC GAT GAT GTG GCT TTG AAG AAC 3′ | 8 | 52% | 27618 |
| X1.1F | 24 | 5′ GAT AGG ATC TTT AGC GAC AGC CGA 3′ | 9 | 50% | 24880 |
| X.1.1R | 24 | 5′ ATG GCG GCC TCT GAG TCC TGG TGG 3′ | 10 | 67% | |
| 2.1F | 24 | 5′ CGG GCT GAA TGC AAT GGA GTG TGC 3′ | 11 | 58% | |
| 3.4F | 18 | 5′ GAC CCC CAT TTG TGT GAC 3′ | 12 | 55.50% | |
| 1060F/S | 19 | 5′ CGA CGA CTC CTG GAG CCC G 3′ | 13 | 73.70% | |
| 1061r/Bio | 24 | 5′ Biotin-TTG ACA CCA GAC CAA CTC GTA ATG 3′ | 14 | 45.80% | |
| 16X.1R | 24 | 5′ AGC CGA CAG CGA TTT CTA GGA TAG 3′ | 15 | 50% | 24879 |
| 17R | 27 | 5′ GTT CTT CAA AGC CAC ATC ATC GCG GTC 3′ | 16 | 52% | 27385 |
| 3′COD R | 28 | 5′ GCT TTC ATT ATC ACT GTC TCC CAG GGT G 3′ | 17 | 50% | 28313 |
| 5′ NCF | 24 | 5′ CAG ACG TTC TTC GCC GAG AGT CGT 3′ | 18 | 58% | 24870 |
| 4869 | 27 | 5′ CAG ACG TTC TTC GCC GAG AGT CGT CGG 3′ | 19 | 63% | |
| NFG | 20 | 5′ CAT TTC GGG GAT TCG GGG GA 3′ | 20 | 60% | |
| NFGP-2 | 20 | 5′ GGG GGA CGG AAC CCG GCG CT 3′ | 21 | 80% | 201880 |
| 767-F | 21 | 5′ CCC TCT ACA CTT ATC ATC TTC 3′ | 22 | 43% | 211616 |
| 16-F | 24 | 5′ CTA TCC TAG AAA TCG CTG TCG GCT 3′ | 23 | 50% | 241173 |
| ECO-F | 24 | 5′ GTC ACT ACT GGA ATT CCC TTC TCC 3′ | 24 | 50% | 24960 |
| ECO-R | 24 | 5′ GGA GAA GGG AAT TCC AGT AGT GAC 3′ | 25 | 50% | 24961 |
| SPF | 21 | 5′ GGA AAT CGC TGT CGC CTA ACC 3′ | 26 | 57% | 211667 |
| SPR | 21 | 5′ GGT AGG CGA CAG CGA TTT CC 3′ | 27 | 57% | 211668 |
| AUAP | 20 | 5′ GGC CAC GCG TCG ACT AGT AC 3′ | 28 | 65% | 202738 |
| NC-F | 20 | 5′ GTA ATG CAC ACTCCA TTG GC 3′ | 29 | 50% | 203814 |
| SNC-F | 18 | 5′ GTA ATG CAC ACT CCA TTG 3′ | 30 | 44% | 181897 |
| BNC-F | 18 | 5′ GCG CTC AGC TGG AAT TCC 3′ | 31 | 55.50% | 181898 |
| BNC-R | 18 | 5′ GGA ATT CCA GCT GAG CGC 3′ | 32 | 61.10% | 181905 |
| pGEX-F | 29 | 5′ GTG GGA TCC CCA TGA CGA CCG CGT CCA CC3′ | 33 | 67% | 29391 |
| pGEX-R1 | 27 | 5′ GAC TCG AGT TAA GCC GAC AGC GAT TTC 3′ | 34 | 51.85% | 27578 |
| pGEX-R2 | 29 | 5′ GAC TCG AGT CAG GGT GAC CGA AAA ATC AG 3′ | 35 | 51.70% | 29396 |
| pGEX-R3 | 31 | 5′ CCCGCTCGAGTCAGGGTGACCGAAAAATCAG 3′ | 36 | 58% | 31277 |

The paragraph beginning at line 6 on page 27 has been amended as follows:

The expression vector (pGEX-5X-1) used for gene fusion construction was the GST Gene Fusion System (Pharmacia). The OFF1 coding region (designated as *"FL"*, full-length) of about 0.5 kb was prepared by PCR with the following 5′ end primer:

5′ GT<u>GGATCC</u>CCATGACGACCGCGTCCA <u>(1-27 of SEQ ID NO:33)</u>, in order
     BamHI

to add a BamHI site 1 base upstream from the start codon ATG and with the 3′ end primer

5′ CCCG <u>CTCGAG</u> TCA GGG TGA CCG AAA AAT CAG 3′ <u>(SEQ ID NO:36) in</u>
     Xho1

order to add an Xho1 site after the stop codon TAA using the PCR kit (Perkin-Elmer/Centus).

5'
TTGACACCAG   ACCAACTGGT   AATGGTAGCG   ACCGGCGCTC   AGCTGGAATT   CCAAAATG

TAATGCACAC   TCCATTGCAT   TCAGCCCGCC   TCTCCTTAGT   CGCCGCCATG   ACGACCGCGT

CCACCTCGCA   GGTGCGCCAG   AACTACCACC   AGGACTCAGA   GGCCGCCATC   AACCGCCAGA

TCAACCTGGA   GCTCTACGCC   TCCTACGTTT   ACCTGTCCAT   GTCTTACTAC   TTTGACCGCG

ATGATGTGGC   TTTGAAGAAC   TTTGCCAAAT   ACTTTCTTCA   CCAATCTCAT   GAGGAGAGGG

AACATGCTGA   GAAACTGATG   AAGCTGCAGA   ACCAACGAGG   TGGCCGAATC   TTCCTTCAGG

ATATCAAGAA   ACCAGACTGT   GATGACTGGG   AGAGCGGGCT   GAATGCAATG   GAGTGTGCAT

TACATTTGGA   AAAAAATGTG   AATCAGTCAC   TACTGGAATT   CCCTTCTCCT   ATCTCTCCCA

GTCCTAGCTG   CTGGCATCAC   TATACTACTA   ACAGACCGCA   ACCTCAACAC   CACCTTCTTC

GACCCCGCCG   GAGGAAGAGA   CCCCATTCTA   TACCAACACC   TATTCTGATT   TTTCGGTCAC

CCTGAAGTTT   ATATTCTTAT   CCTACCAGGC   TTCGGAATAA   TCTCCCATAT   TGTAACTTAC

TACTCCGGAA   ATCGCTGTCG   CCTAACCGCT   AACATTACTG   CAGGCCACCT   ACTCATGCAC

CTAATTGGAA   GCGCCACCCT   AGCAATATCA   ACCATTAACC   TTCCCTCTAC   ACTTATCATC

TTCACAATTC   TAATTCTACT   GACTATCCTA   GAAATCGCTG   TCGCCTTAAT   CCAAGCCTAC

GTTTTCACAC   TTCTAGTAAG   CCTCTACCTG   CACGACAACA   CATAAAAAAA   AA   3'

Fig. 1

Fig. 2A

CLONE p47　　　GGGGGACGGAACCCGG

CGCTCGTTCCCCACCCCGGCCGGCCGCCCATAGCCAGCCCTCCGTCAC

CLONE T 16　　　　　　　　　TTGACACCXX

CTCTTCACCGCACCCTCGGACTGCCCCAAGGCCCCCGCCGCCGCTCC
AGACCAACTGGTAATGGTAGCGACCGGCGCTCAGCTGGAATTCCAAAA

AGCGCCGCGCAGCCACCGCCGCCGCCGCCGCCTCTCCTTAGTCGCCGCC
AATGTAATGCACACTCCATTGCATTCAGCCCGCCTCTCCTTAGTCGCCGCC

| ATG | ACG | ACC | GCG | TCC | ACC | TCG | CAG | GTG | CGC | CAG |
| ATG | ACG | ACC | GCG | TCC | ACC | TCG | CAG | GTG | CGC | CAG |

| AAC | TAC | CAC | CAG | GAC | TCA | GAG | GCC | GCC | ATC | AAC |
| AAC | TAC | CAC | CAG | GAC | TCA | GAG | GCC | GCC | ATC | AAC |

| CGC | CAG | ATC | AAC | CTG | GAG | CTC | TAC | GCC | TCC | TAC |
| CGC | CAG | ATC | AAC | CTG | GAG | CTC | TAC | GCC | TCC | TAC |

| GTT | TAC | CTG | TCC | ATG | TCT | TAC | TAC | TTT | GAC | CGC |
| GTT | TAC | CTG | TCC | ATG | TCT | TAC | TAC | TTT | GAC | CGC |

| GAT | GAT | GTG | GCT | TTG | AAG | AAC | TTT | GCC | AAA | TAC |
| GAT | GAT | GTG | GCT | TTG | AAG | AAC | TTT | GCC | AAA | TAC |

| TTT | CTT | CAC | CAA | TCT | CAT | GAG | GAG | AGG | GAA | CAT |
| TTT | CTT | CAC | CAA | TCT | CAT | GAG | GAG | AGG | GAA | CAT |

| GCT | GAG | AAA | CTG | ATG | AAG | CTG | CAG | AAC | CAA | CGA |
| GCT | GAG | AAA | CTG | ATG | AAG | CTG | CAG | AAC | CAA | CGA |

| GGT | GGC | CGA | ATC | TTC | CTT | CAG | GAT | ATC | AAG | AAA |
| GGT | GGC | CGA | ATC | TTC | CTT | CAG | GAT | ATC | AAG | AAA |

| CCA | GAC | TGT | GAT | GAC | TGG | GAG | AGC | GGG | CTG | AAT |
| CCA | GAC | TGT | GAT | GAC | TGG | GAG | AGC | GGG | CTG | AAT |

| GCA | ATG | GAG | TGT | GCA | TTA | CAT | TTG | GAA | AAA | AAT |
| GCA | ATG | GAG | TGT | GCA | TTA | CAT | TTG | GAA | AAA | AAT |

| GTG | AAT | CAG | TCA | CTA | CTG | GAA | CTG | CAC | AAA | CTG |
| GTG | AAT | CAG | TCA | CTA | CTG | GAA | TTC | CCT | TCT | CCT |

| GCC | ACT | GAC | AAA | AAT | GAC | CCC | CAT | TTG | TGT | GAC |
| ATC | TCT | CCC | AGT | CCT | AGC | TGC | TGG | CAT | CAC | TAT |

SUBSTITUTE SHEET (RULE 26)

| TTC | ATT | GAG | ACA | CAT | TAC | CTG | AAT | GAG | CAG | GTG |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ACT | ACT | AAC | AGA | CCG | CAA | CCT | CAA | CAC | CAC | CTT |

| AAA | GCC | ATC | AAA | GAA | TTG | GGT | GAC | CAC | GTG | ACC |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| CTT | CGA | CCC | CGC | CGG | AGG | AAG | AGA | CCC | CAT | TCT |

| AAC | TTG | CGC | AAG | ATG | GGA | GCG | CCC | GAA | TCT | GGC |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ATA | CCA | ACA | CCT | ATT | CTG | ATT | TTT | CGG | TCA | CCC |

| TTG | GCG | GAA | TAT | CTC | TTT | GAC | AAG | CAC | ACC | CTG |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

TGA AGTTTATATTCTTATCCTACCAGGCTTCGGAATAATCTCCCATATT

| GGA | GAC | AGT | GAT | AAT | GAA | AGC | TAA |
|-----|-----|-----|-----|-----|-----|-----|-----|

GCCTCGGGCTAATT
GTAACTTACTACTCCGGAAATCGCTGTCGCCTAACCGCTAACATTACTGC

TCCCATAGCCGTGGGGTGACTTCCCTGGTCACCAAGGCAGTGCATGCAT
AGGCCACCTACTCATGCACCTAATTGGAAGCGCCACCCTAGCAATATCA

GCATGTTGGGGTTTCCTTTACCTTTTCTATAAGTTGTACCAAAACATCCAC
ACCATTAACCTTCCCTCTACACTTATCATCTTCACAATTCTAATTCTACTG

TTAAGTTCTTTGATTTGTACCATTCCTTCAAATAAAGAAATTTGGTACCCA
ACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTACGTTTTCACACT

AAAAAAAA                              C
TCTAGTAAGCCTCTACCTGCACGACAAXACATAAAAA AAA

Fig. 2A Cont.

TTGACACCAGACCAACTGGTAATGGTAGCGACCGGCGCTCAGCTGGAATTCCAAAAATGT

AATGCACACTCCATTGCATTCAGCCCGCCTCTCCTTAGTCGCCGCC

| met | thr | thr | ala | ser | thr | ser | gln | val | arg | gln |
| ATG | ACG | ACC | GCG | TCC | ACC | TCG | CAG | GTG | CGC | CAG |

| asn | tyr | his | gln | asp | ser | glu | ala | ala | ile | asn |
| AAC | TAC | CAC | CAG | GAC | TCA | GAG | GCC | GCC | ATC | AAC |

| arg | gln | ile | asn | leu | glu | leu | tyr | ala | ser | tyr |
| CGC | CAG | ATC | AAC | CTG | GAG | CTC | TAC | GCC | TCC | TAC |

| val | tyr | leu | ser | met | ser | tyr | tyr | phe | asp | arg |
| GTT | TAC | CTG | TCC | ATG | TCT | TAC | TAC | TTT | GAC | CGC |

| asp | asp | val | ala | leu | lys | asn | phe | ala | lys | tyr |
| GAT | GAT | GTG | GCT | TTG | AAG | AAC | TTT | GCC | AAA | TAC |

| phe | leu | his | gln | ser | his | glu | glu | arg | gln | his |
| TTT | CTT | CAC | CAA | TCT | CAT | GAG | GAG | AGG | GAA | CAT |

| ala | glu | lys | leu | met | lys | leu | gln | asn | gln | arg |
| GCT | GAG | AAA | CTG | ATG | AAG | CTG | CAG | AAC | CAA | CGA |

| gly | gly | arg | ile | phe | leu | gln | asp | ile | lys | lys |
| GGT | GGC | CGA | ATC | TTC | CTT | CAG | GAT | ATC | AAG | AAA |

| pro | asp | cys | asp | asp | trp | glu | ser | gly | leu | asn |
| CCA | GAC | TGT | GAT | GAC | TGG | GAG | AGC | GGG | CTG | AAT |

| ala | met | glu | cys | ala | leu | his | leu | glu | lys | asn |
| GCA | ATG | GAG | TGT | GCA | TTA | CAT | TTG | GAA | AAA | AAT |

| val | asn | gln | ser | leu | leu | glu | phe | pro | ser | pro |
| GTG | AAT | CAG | TCA | CTA | CTG | GAA | TTC | CCT | TCT | CCT |

| ile | ser | pro | ser | pro | ser | cys | trp | his | his | thr |
| ATC | TCT | CCC | AGT | CCT | AGC | TGC | TGG | CAT | CAC | TAT |

| thr | thr | asn | arg | pro | glu | pro | gln | his | his | leu |
| ACT | ACT | AAC | AGA | CCG | CAA | CCT | CAA | CAC | CAC | CTT |

| leu | arg | pro | arg | arg | arg | lys | arg | pro | his | ser |
| CTT | CGA | CCC | CGC | CGG | AGG | AAG | AGA | CCC | CAT | TCT |

| ile | pro | thr | pro | ile | leu | ile | phe | arg | ser | pro |
| ATA | CCA | ACA | CCT | ATT | CTG | ATT | TTT | CGG | TCA | CCC |

TGA AGTTTATATTCTTATCCTACCAGGCTTCGGAATAATCTCCCATATTGTAACTTAC

TACTCCGGAAATCGCTGTCGCCTAACCGCTAACATTACTGCAGGCCACCTACTCATGCAC

CTAATTGGAAGCGCCACCCTAGCAATATCAACCATTAACCTTCCCTCTACACTTATCATC

TTCACAATTCTAATTCTACTGACTATCCTAGAAATCGCTGTCGCCTTAATCCAAGCCTAC

GTTTTCACACTT̶A̶T̶A̶T̶A̶G̶C̶C̶T̶A̶A̶A̶A̶A̶A̶A̶A̶
CTAGTAAGCCTCTACCTGCACGACAACACATAAAAAAA

**Fig. 5**

|  | | | BNC | | |
|---|---|---|---|---|---|
| 1061<br>TTGACACCAG | ACCAACTGGT | AATGGTAGCG | ACCGGCGCTC | AGCTGGAATT | CGAAAAAATG |
| NCS<br>TAATGCACAC | TCCATTGCAT | TCAGCCCGCC | TCTCCTTAGT | CGCCGCCATG | ACGACCGCGT |
| CCACCTCGCA | GGTGCGCCAG | AACTACCACC | X1<br>AGGACTCAGA | GGCCGCCATG | AACCGCCAGA |
| TCAACCTGGA | GCTCTACGCC | TCCTACGTTT | ACCTGTCCAT | GTCTTACTAC | 17<br>TTTGACCGCG |
| 17<br>ATGATGTGGC | TTTGAAGAAC | TTTGCCAAAT | ACTTTCTTCA | CCAATCTCAT | GAGGAGAGGG |
| AACATGCTGA | GAAACTGATG | AAGCTGCAGA | ACCAACGAGG | TGGCCGAATC | TTCCTTCAGG |
| ATATCAAGAA | ACCAGACTGT | GATGACTGGG | AGAGCGGGCT | 2.1<br>GAATGCAATG | GAGTGTGCAT |
| TACATTTGGA | AAAAAATGTG | AATCAGTCAG | ECOF<br>TACTGGAATT | CCCTTCTCCT | ATCTCTCCCA |
| GTCCTAGCTG | CTGGCATCAC | TATACTACTA | ACAGACCGCA | ACCTCAACAC | CACCTTCTTC |
| GACCCCGCCG | GAGGAAGAGA | CCCCATTCTA | TACCAACACC | TATTCTGATT | TTTCGGTCAC |
| CCTGAAGTTT | ATATTCTTAT | CCTACCAGGC | TTCGGAATAA | TCTCCCATAT | TGTAACTTAC |
| TACTCCGGAA | SPF<br>ATCGCTGTCG | CCTAACCGCT | AACATTACTG | CAGGCCACCT | ACTCATGCAC |
| CTAATTGGAA | 723<br>GCGCCACCCT | AGCAATATCA | ACCATTAACC | TTCCCTCTAC | 767<br>ACTTATCATC |
| 767<br>TTCACAATTC | TAATTCTACT | GACTATCCTA | 16<br>GAAATCGCTG | TCGCCTTAAT | CCAAGCCTAC |
| GTTTTCACAC | TTCTAGTAAG | CCCTCTACCTG | CCACGACAACA | CCATAAAAAAA | AA |

Fig. 7

TTGACACCAG   ACCAACTGGT   AATGGTAGCG   ACCGGCGCTC   AGCTGGAATT   CCAAAAAATG

TAATGCACAC   TCCATTGCAT   TCAGCCCGCC   TCTCCTTAGT   CGCCGCCAGG   ACGACCGCGT

CCACCTCGCA   GGTGCGCCAG   AACTACCACC   AGGACTCAGA   CGCCGCCATC   AACCGCCAGA

TCAACCTGGA   GCTCTACGCC   TCCTACGTTT   ACCTGTCCAT   GTCTTACTAC   TTTGACCGCG

ATGATGTGGC   TTTGAAGAAC   TTTGCCAAAT   ACTTTCTTCA   CCAATCTCAT   GAGGAGAGGG

                                 PstI

AACATGCTGA   GAAACTGATG   AAGCTGCAGA   ACCAACGAGG   TGGCCGAATC   TTCCTTCAGG

ATATCAAGAA   ACCAGACTGT   GATGACTGGG   AGAGCGGGCT   GAATGCAATG   GAGTGTGCAT

                                 ECOR1

TACATTTGGA   AAAAAATGTG   AATCAGTCAC   TACTGGAATT   CCCTTCTCCT   ATCTCTCCCA

GTCCTAGCTG   CTGGCATCAC   TATACTACTA   ACAGACCGCA   ACCTCAACAC   CACCTTCTTC

GACCCCGCCG   GAGGAAGAGA   CCCCATTCTA   TACCAACACC   TATTCTGATT   TTTCGGTCAC

CCTCGAGTTT   ATATTCTTAT   CCTACCAGGC   TTCGGAATAA   TCTCCCATAT   TGTAACTTAC

TACTCCGGAA   ATCGCTGTCG   CCTAACCGCT   AACATTACTG   CAGGCCACCT   ACTCATGCAC

CTAATTGGAA   GCGCCACCCT   AGCAATATCA   ACCATTAACC   TTCCCTCTAC   ACTTATCATC

TTCACAATTC   TAATTCTACT   GACTATCCTA   GAAATCGCTG   TCGCCTTAAT   CCAAGCCTAC

GTTTTCACAC   TTCTAGTAAG   CCCTCTACCTG   CCACGACAACA   CCATAAAAAAA   AA

# Fig. 8